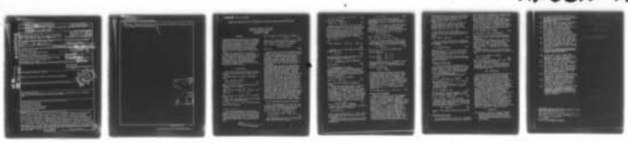
AD-A050 305 STANFORD UNIV CALIF DEPT OF ELECTRICAL ENGINEERING F/G 9/4
SQUARE-ROOT ALGORITHMS FOR THE CONTINUOUS-TIME LINEAR LEAST SQU--ETC(U)
DEC 77 M MORF, B LEVY, T KAILATH F44620-74-C-0068
UNCLASSIFIED AFOSR-TR-78-0146 NL

1 OF 1
AD
A050 305



END
DATE
FILMED
3-78
DDC

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD A050305

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 78 - 0146	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SQUARE-ROOT ALGORITHMS for the Continuous-Time LINEAR LEAST SQUARES ESTIMATION PROBLEM.		5. TYPE OF REPORT & PERIOD COVERED Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) M. Morf, B. Lévy [redacted] T. Kailath		8. CONTRACT OR GRANT NUMBER(s) ~~F44-620-74-C-0068~~ F44620-74-C-0068, N00014-75-C-0601
9. PERFORMING ORGANIZATION NAME AND ADDRESS Stanford University Dept. of Electrical Engineering Stanford, CA. 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A6
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scienfitic Research (NM) Bolling Air Force Base, Bldg. 410 Washington, D.C. 20332		12. REPORT DATE Dec[redacted]77
		13. NUMBER OF PAGES 6 p.
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Proc. 1977 IEEE Conference on Decision & Control, New Orleans, LA., December 1977.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Triangular Square-Root
Linear State Estimator
Chandrasekhar-type equations
Fast Square-Root Algorithms
Stochastic Calculus

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A simple differential equation for the triangular square-root of the error covariance of the linear state estimator is derived. Previous algorithms involved an antisymmetric matrix in the square-root differential equation. In the constant model case, Chandrasekhar-type equations are shown to constitute a set of fast square-root algorithms for the derivative of the error variance.

Square-Root algorithms for the smoothing problem are presented and as in the discrete case, an array method for handling continuous square-roots is developed. This method also yields very naturally the usual normalizations of stochastic

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

400852

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20.

calculus, suggesting extensions to more general stochastic equations, even to estimators for nonlinear models.

AFOSR-TR- 78-0146

SQUARE-ROOT ALGORITHMS FOR THE CONTINUOUS-TIME LINEAR LEAST SQUARES ESTIMATION PROBLEM†

M. Morf, B. Lévy and T. Kailath
Information Systems Laboratory
Stanford University
Stanford, CA 94305

Abstract

A simple differential equation for the triangular square-root of the error covariance of the linear state estimator is derived. Previous algorithms involved an antisymmetric matrix in the square-root differential equation. In the constant model case, Chandrasekhar-type equations are shown to constitute a set of fast square-root algorithms for the derivative of the error variance.

Square-root algorithms for the smoothing problem are presented and as in the discrete case, an array method for handling continuous square-roots is developed. This method also yields very naturally the usual normalizations of stochastic calculus, suggesting extensions to more general stochastic equations, even to estimators for nonlinear models.

I. Introduction

Since a full version of this paper will appear in [1], we present here only an outline of our *results*.

We assume that we are given a state-space model

$$\dot{x}(t) = F(t)x(t) + G(t)u(t)$$

$$y(t) = H(t)x(t) + v(t) \qquad (1a)$$

where $u(\cdot)$ and $v(\cdot)$ are white noises with normalized joint covariance:

$$E\left\{\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} [u^T(s)\ v^T(s)]\right\} = \begin{bmatrix} I_q & 0 \\ 0 & I_p \end{bmatrix} \delta(t-s) \qquad (1b)$$

Here, T denotes the transpose and E the expectation.

Then, $\hat{x}(t)$, the linear least squares estimate of $x(t)$ given the observed data $Y_o^t \overset{\Delta}{=} \{y(s)\ ,\ 0 \leq s \leq t\}$ can be obtained via the Kalman filter equation

$$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + P(t)H^T(t)(y(t) - H(t)\hat{x}(t))\ ,$$
$$\hat{x}(0) = 0 \qquad (2a)$$

where $P(t)$, the covariance of the error of the state estimate, obeys the Riccati equation

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)G^T(t) - P(t)H^T(t)H(t)P(t)$$

$$P(0) = \Pi_o \qquad (2b)$$

Over the years, the Riccati equation has been studied in great detail, and as one way to improve its numerical conditioning, a whole family of alternative square-root algorithms has been introduced. In the discrete-case, square-root algorithms have been studied by Potter [2], Golub [3], Schmidt [4], Kaminski and Bryson [5], [6], Bierman [7], [8], Morf and Kailath [9], among others.

Continuous-time square-root algorithms have attracted somewhat less interest (see Andrews [10], Tapley and Choe [11], Bierman [12], and Morf and Kailath [9], Appendix C). Somewhat surprisingly, all known solutions explicitly introduce a certain antisymmetric matrix into the differential equation for the square-roots. In Section II of this paper, we present a differential equation for the triangular square-roots that does not explicitly contain any such antisymmetric matrix. We show that these matrices are generators of the orthogonal transformations that relate the various square-roots and that a particular choice of this matrix will result in triangular square-roots. Among other results, we shall also discuss the stability of our new equations, and we shall also present some alternative forms (such as the information filter forms for the case of high initial uncertainty of the state estimate). In passing we also note that these ideas can also be applied to the Chandrasekhar-type algorithms (for constant models).

II. Continuous-time square-root algorithms

If S is a square-root of P and $P(\cdot)$ is strictly positive definite, so that we can write $P = SS^T$ where S is nonsingular, then the Riccati equation (2b) can be rewritten as

$$\dot{P} = \dot{S}\, S^T + S\, \dot{S}^T = (F - \tfrac{1}{2} S\, S^T H^T H) S\, S^T + \tfrac{1}{2} G\, G^T\, S^{-T} S^T + S\, S^T (F^T - \tfrac{1}{2} H^T H\, S\, S^T) + \tfrac{1}{2} S\, S^{-1}\, G\, G^T\ . \qquad (3)$$

This equation is satisfied if the square-root S obeys the differential equation

---

† This work was supported by the Air Force Office of Scientific Research, Air Force Systems Command, under Contract AF44-620-74-C-0068, and in part by the National Science Foundation under Contract NSF-Eng-75-18952 and the Joint Services Electronics Program under Contract N00014-75-C-0601.

$$\dot S = (F - \tfrac{1}{2} S S^T H^T H)S + \tfrac{1}{2} G G^T S^{-T} \qquad (4a)$$

with initial conditions

$$S(0) = S_o, \quad S_o S_o^T = \Pi_o \qquad (4b)$$

However, it is well-known that matrix square-roots are not unique, for if S is a square-root, so is ST, where T is any orthonormal matrix. To study how this non-uniqueness is reflected into the differential equation, we note first that any differentiable orthonormal matrix $T(t)$ can be uniquely characterized by a skew-or-anti-symmetric matrix generator $\mathcal{A}(\cdot)$ such that

$$\mathcal{A}(\cdot) + \mathcal{A}^T(\cdot) = 0 \qquad (5a)$$

and

$$\dot T(t) = A(t)T(t), \quad T(0) = T_o \quad T_o T_o^T = I_n. \qquad (5b)$$

Therefore, if S is the square-root obeying equation (4) and if

$$S_{\mathcal{a}} = ST \qquad (6a)$$

by differentiation of (6a), we obtain the differential equation satisfied by $S_{\mathcal{a}}$

$$\dot S_{\mathcal{a}} = (F - \tfrac{1}{2} S_{\mathcal{a}} S_{\mathcal{a}}^T H^T H)S_{\mathcal{a}} + (S_{\mathcal{a}} \mathcal{A} S_{\mathcal{a}}^T + \tfrac{1}{2} GG^T)S_{\mathcal{a}}^{-T},$$

$$S_{\mathcal{a}}(0) = S_o T_o. \qquad (6b)$$

In this equation, the matrix $A(t) \triangleq S(t)\mathcal{A}(t)S^T(t)$ is antisymmetric and can be made arbitrary since $\mathcal{A}(t)$ is arbitrary. Therefore, this establishes that all the square-roots of P are given by the solutions of the differential equation

$$\dot S = (F - \tfrac{1}{2} SS^T H^T H)S + (A + \tfrac{1}{2} GG^T)S^{-T}$$

$$S(0) = S_o, \quad S_o S_o^T = \Pi_o \qquad (7)$$

where $A(\cdot)$ is an arbitrary antisymmetric matrix.

Clearly, the square-root matrix S solution of (7) depends of the antisymmetric matrix $A(\cdot)$ chosen. This dependence gives us an additional degree of freedom, which can be exploited in different ways. For example, a difficulty with formula (7) is that S^{-1} has to be computed at each integration step. However, Andrews [10] noted that $A(\cdot)$ could be chosen so as to make S lower triangular, so that S^{-1} can be computed more easily. To do this, rewrite (7) as

$$\dot S = (A + \Gamma)S^{-T} \qquad (8)$$

with $\Gamma \triangleq F\,SS^T + \tfrac{1}{2}(GG^T - SS^T H^T HSS^T)$.

Then, S can be maintained in lower triangular form by calculating A such that

$$[(A + \Gamma)S^{-T}]_- = 0$$

and by inserting this value of A in (8). (Here $[M]_-$ denotes the strictly upper triangular part of the matrix M).

A somewhat more direct approach was discovered by Tapley and Choe [11] who noted that (8), when rewritten as

$$A + \Gamma = \dot S S^T \qquad (10)$$

was giving a lower-times upper-triangular factorization of $A + \Gamma$, and thereby were able to simultaneously compute $\dot S$ (in lower triangular form) and A.

Actually, it is possible to find a differential equation for S in lower triangular form that does not explicitly involve any antisymmetric matrix in the equation: multiply (3) on the left by S^{-1} and on the right by S^{-T} to obtain

$$S^{-1}\dot S + \dot S^T S^{-T} = M$$

where $M \triangleq \bar F + \bar F^T + \bar G \bar G^T - \bar H^T \bar H$

and $\bar F \triangleq S^{-1}FS$, $\bar G = S^{-1}G$, $\bar H = HS$.

Since S is lower triangular, $S^{-1}\dot S$ is the lower triangular part of M and we can identify (an idea analogous to the Wiener-Hopf technique)

$$\dot S = S[M]_{+/2} \qquad (11)$$

Here, $[\ ]_{+/2}$ is the "lower-triangular part" operator defined by

$$[U]_{+/2ij} = \begin{cases} 0 & \text{for } i < j \\ \tfrac{1}{2}u_{ii} & \text{for } i = j \\ u_{ij} & \text{for } i < j \end{cases}$$

for an arbitrary matrix U. The "strictly lower-triangular part" has zero values for $i = j$. Similarly, $[\ ]_{-/2}$, the "upper-triangular part" operator is defined by

$$[U]_{-/2} = [U^T]^T_{+/2} .$$

The main advantage of equation (11) over those of Andrews and Tapley and Choe is that it does not involve explicitly any antisymmetric matrix in the differential equation. Moreover, the matrices $(\bar F, \bar G, \bar H)$ have a nice interpretation: these matrices arise in the dynamic model for the modified state-variable $n = S^{-1}x$. The significance of this choice of variables is that the variance of the error $\tilde n$ will be I. The dynamic model for n can be obtained from (1) and (11) as

$$\dot n = (\bar F - [M]_{+/2})n + \bar G u, \quad y = \bar H n + v \qquad (12)$$

<u>Stability of the equation</u>

Square-root methods have several advantages over the Riccati equation (2b): for example, $P = SS^T$ is always non-negative definite, a property which is not always guaranteed by direct numerical integration of (26). Furthermore, since the "condition number" of S is the square-root of the one for P, P can be computed with greater accuracy. On the other hand, these algorithms require a slightly larger amount of computations, as is shown in Appendix B of [1] where computational aspects are discussed.

It should also be pointed out that the stability properties of the Riccati equation are conserved by equation (11): as in [14], assume that the pair (F,G) (respectively (H,F)) is uniformly completely controllable (respectively, observable) and that S_1 and S_2 are two solutions of the square-root equation (11) corresponding to different initial conditions S_{10} and S_{20} such that $S_{10}S_{10}^T = \Pi_{10} > 0$ and $S_{20}S_{20}^T = \Pi_{20} > 0$.

In this case, $P_1 = S_1S_1^T$ and $P_2 = S_2S_2^T$ are positive-definite solutions of the Riccati equation (2b) and if $\Delta P = P_1 - P_2$, it was proved by Kalman in [14] that $\Delta P(t) \to 0$ as $t \to \infty$.

Now, denote $\Delta S = S_1 - S_2$, then $\Delta P = \Delta S \,.\, S_2^T + S_1 \,.\, \Delta S^T$ and by multiplication on the left by S_1^{-1} and on the right by S_2^{-T}, we obtain

$$S_1^{-1}\Delta S + \Delta S^T S^{-T} = S_1^{-1}\Delta P\, S_2^{-T} \quad . \tag{13}$$

But S_1 and S_2 are both lower triangular, and therefore $S_1^{-1}\Delta S$ is the lower triangular part of $S_1^{-1}\Delta P_2^{-T}$:

$$\Delta S = S_1[S_1^{-1}\Delta P\, S_2^{-T}]_{+/2} \tag{14}$$

and since $\Delta P(t) \to 0$ as $t \to \infty$ and S_i and S_i^{-1} are bounded for $i = 1,2$ (P_i and P_i^{-1} are proved to be bounded in [14] we can conclude that $\Delta S(t) \to 0$ as $t \to \infty$.

This indicates that the square-root equation (11) and the Riccati equation (2b) have the same stability properties.

Information filter forms

In several instances, for example when Π_o (the uncertainty on the initial condition $x(0)$) becomes large, it is more convenient to compute $P^{-1}(t)$ instead of $P(t)$. Filters involving P^{-1} or its square-root $\Omega(P^{-1} = \Omega\,\Omega^T)$ have been called information filters.

In this case, Ω obeys the differential equation

$$\dot{\Omega} = (-F^T - \tfrac{1}{2}\,\Omega\Omega^T\, GG^T)\Omega + (\alpha + \tfrac{1}{2}H^TH)\Omega^{-T} \tag{15}$$

where $\Omega(0)\Omega^T(0) = \Pi_o^{-1}$ and α is an arbitrary antisymmetric matrix.

Then, instead of computing $\hat{x}$, it is more convenient to compute $\hat{d} = P^{-1}\hat{x}$ with the information filter

$$\dot{\hat{d}} = -(F + GG^T\Omega\Omega^T)^T\,\hat{d} + H^Ty \quad . \tag{16}$$

Alternatively, we can also propagate $\hat{n} = \Omega^T\hat{x}$ via the variance normalized information filter

$$\dot{\hat{n}} = (\bar{\alpha} - \tfrac{1}{2}\,\bar{G}\,\bar{G}^T - \tfrac{1}{2}\,\bar{H}^T\bar{H})\hat{n} + \bar{H}^Ty \tag{17}$$

where $\bar{\alpha} \triangleq \Omega^{-1}\alpha\Omega^{-T}$, $\bar{G} = \Omega^TG$, $\bar{H} \triangleq H\Omega^{-T}$,

We have seen in (12) that such a filter arises naturally in the interpretation of equation (11) for the triangular square-root S. The information filter form of equation (11) is

$$\dot{\Omega}^T = -[M]_{+/2}\,\Omega^T \tag{18}$$

where Ω is the upper triangular square-root of P^{-1}.

Time-Invariant Systems

In the constant model case (H, F, G constant), it is known [13] that fast square-root algorithms can be obtained by propagating the square-root of $\dot{P}(t)$ instead of the square-root of $P(t)$. However our earlier discussion suggests that we should consider a family of square roots, differing from each other by orthogonal transformations.

To see how this can be incorporated let us review the derivation of the Chandrasekhar equations. The first step is to exploit the constancy of the model parameters by differentiating the Riccati equation (2b) to obtain

$$\ddot{P} = (F - K(t)H)\dot{P} + \dot{P}(F - K(t)H)^T \tag{19}$$

and therefore $\dot{P}(t) = \phi_k(t,0)\dot{P}(0)\phi_k^T(t,0)$

where $\phi_k(t,0)$ is the transition matrix associated to $(F - K(t)H)$, $K(t) = P(t)H^T$ being the Kalman gain.

Now, consider for simplicity the special case where $\Pi_o = 0$ (known initial conditions). In this case $\dot{P}(0) = GG^T$, so that we can factor $\dot{P}(t) = L(t)L^T(t)$ in square-root form, $L(t)$ being $n\times\alpha$ with $\alpha = \text{rank } GG^T \leq q$: the number of inputs.

Then substituting $\dot{P} = LL^T$ in (19), we find that $L(t)$ obeys the Chandrasekhar-type equations [13]

$$\begin{aligned} \dot{L}(t) &= (F - K(t)H)L(t) - L(t)a(t) \\ \dot{K}(t) &= L(t)L^T(t)H^T \end{aligned} \tag{20}$$

where $L(0)L^T(0) = GG^T$, $K(0) = \Pi_oH^T$, $a(t)$ being some $\alpha\times\alpha$ antisymmetric matrix.

In [12], $a(t) = 0$ was chosen, however the introduction of $a(t)$ in (20) can provide some additional control over the numerical behavior of the square-root $L(t)$.

III. Conclusions

In [1], we show that the results of Section II can be extended to the fixed-point smoothing problem. Also, continuous-time array methods similar to those developed in [9] by Morf and Kailath for the discrete-time case are presented. This approach yields very naturally the usual normalizations of stochastic calculus, suggesting extensions to more general stochastic equations.

In Appendix A of [1], we discuss an alternative set of variance normalized Chandrasekhar-type equations that involve only one differential equation for the square-root of the error covariance; however, the variance normalized state model parameters have to be obtained. In Appendix B, we present operation counts for the algorithm of Section II and those of Andrews and Tapley and Choe. The square-root algorithms involve roughly 15 to 30% more computations than the Riccati equation. Our new equation has the advantage of giving an explicit differential equation for the square-root, which is consequently easier to implement for analog simulation.

References

[1] M. Morf, B. Lévy and T. Kailath, "Square-Root Algorithms for the Continuous-Time Linear Least Squares Estimation Problem," to appear in the IEEE Transactions on Automatic Control.

[2] J.E. Potter and R.G. Stern, "Statistical filtering of space navigation measurements", Proc 1963 AIAA Guidance and Control Conference, 1963.

[3] G.H. Golub, "Numerical Methods for Solving Linear Least-Squares Problems", Numer. Math., vol. 7, pp. 206-216, 1965.

[4] S.F. Schmidt, "Computational Techniques in Kalman Filtering", NATO AGARD-139, Ed. Leondes, pp. 67-103, Feb. 1970.

[5] P.G. Kaminski, "Square-root filtering and smoothing for discrete processes", Ph.D. Dissertation, Dep. Aerosp. Astronaut. Stanford Univ., Stanford, CA., Sept. 1971.

[6] P.G. Kaminski and A.E. Bryson, "Discrete square-root smoothing", Proc. AIAA Guidance and Control Conf., paper No. 72-877, 1972.

[7] G.J. Bierman, "Sequential Square-Root Filtering and Smoothing of Discrete Linear Systems", Automatica, vol. 10, pp. 147-158, March 1974.

[8] G.J. Bierman, Factorization Algorithms for Discrete Sequential Estimation, New York: Academic Press, March 1977.

[9] M. Morf and T. Kailath, "Square-Root Algorithms for Least-Squares Estimation", IEEE Trans. on Automatic Control, Vol. AC-20, No. 4, pp. 487-497, Aug. 1975.

[10] A. Andrews, "A square root formulation of the Kalman covariance equations", AIAA J., vol. 6, pp. 122-123, Feb. 1976.

[11] B.D. Tapley and C.Y. Choe, "An Algorithm for Propagating the Square-Root Covariance Matrix in Triangular Form", IEEE Transactions on Automatic Control, Vol. AC-21, No. 1, pp. 122-123, Feb. 1976.

[12] G.J. Bierman, "A Square-Root Data Array Solution of the Continuous-Discrete Filtering Problem", IEEE Trans. on Automatic Control, Vol. AC-18, No. 6, pp. 675-676.

[13] T. Kailath, "Some New Algorithms for recursive estimation in constant linear systems", IEEE Trans. Inform. Theory, Vol. IT-19, pp. 750-760.

[14] R.E. Kalman, "New Methods in Wiener Filtering Theory", Proceedings of the First Symposium on Engineering Applications of Random Function Theory, J. Bogdanoff and F. Kozin editors, John Wiley: New York, 1963, pp. 270-388.